O-GlcNAcase Fragment Discovery with Fluorescence Polarimetry

Borodkin, Vladimir S.; Rafie, Karim; Selvan, Nithya; Aristotelous, Tonia; Navratilova, Iva; Ferenbach, Andrew T.

Published in:
ACS Chemical Biology

DOI:
10.1021/acschembio.8b00183

Publication date:
2018

Document Version
Peer reviewed version

O-GlcNAcase fragment discovery with fluorescence polarimetry

Journal:	*ACS Chemical Biology*
Manuscript ID	cb-2018-00183u.R2
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	BORODKIN, Vladimir; University of Dundee, Gene Rgulation and Expression Rafie, Karim; University of Dundee, Centre for Gene Regulation and Expression, School of Life Sciences Selvan, Nithya; University of Georgia Complex Carbohydrate Research Center Aristotelous, Tonia; University of Dundee College of Life Sciences Navratilova, Iva; University of Dundee, College of Life Sciences Ferenbach, Andrew; University of Dundee, Centre for Gene Regulation and Expression, School of Life Sciences van Aalten, Daan; University of Dundee, Centre for Gene Regulation and Expression, School of Life Sciences

This is the accepted manuscript version of the following article:
Borodkin, VS, Rafie, K, Selvan, N, Aristotelous, T, Navratilova, I, Ferenbach, AT & Van Aalten, DMF 2018, '**O-GlcNAcase Fragment Discovery with Fluorescence Polarimetry**' ACS Chemical Biology, vol 13, no. 5, pp. 1353-1360. DOI: 10.1021/acschembio.8b00183

O-GlcNAcase fragment discovery with fluorescence polarimetry

Vladimir S. Borodkin[1,*#], Karim Rafie[1#], Nithya Selvan[1,3#], Tonia Aristotelous[2], Iva Navratilova[2], Andrew T. Ferenbach[1] and Daan M. F. van Aalten[1*]

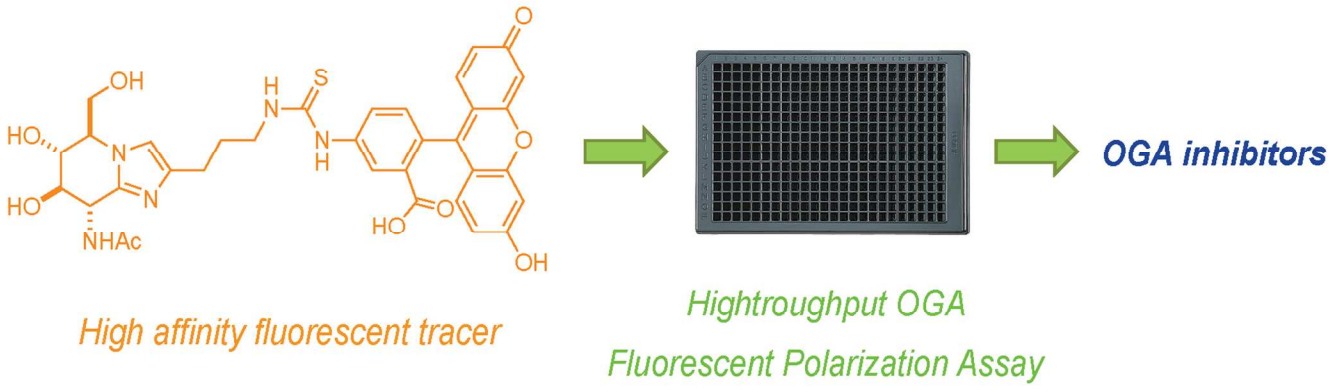
[1]Centre for Gene Regulation and Expression and [2]Division of Biological Chemistry and Drug Discovery, School of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH, United Kingdom. [3]Present address: Complex Carbohydrate Research Center, University of Georgia, Athens, GA, United States.

[*]To whom correspondence should be addressed at vsborodkin@dundee.ac.uk or dmfvanaalten@dundee.ac.uk

[#]These authors contributed equally

Abstract

The attachment of the sugar D-N-acetylglucosamine (GlcNAc) with to specific serine and threonine residues on proteins is referred to as protein O-GlcNAcylation. O-GlcNAc transferase (OGT) is the enzyme responsible for carrying out the modification while O-GlcNAcase (OGA) reverses it. Protein O-GlcNAcylation has been implicated in a wide range of cellular processes including transcription, proteostasis and stress response. Dysregulation of O-GlcNAc has been linked to diabetes, cancer, neurodegenerative and cardiovascular disease. OGA has been proposed to be a drug target for the treatment of Alzheimer's and cardiovascular disease given that increased O-GlcNAc levels appear to exert a protective effect. The search for specific, potent and drug-like OGA inhibitors with bioavailability in the brain is therefore a field of active research, requiring orthogonal high-throughput assay platforms. Here we describe the synthesis of a novel probe for use in a fluorescence polarization based assay for the discovery of inhibitors of OGA. We show that the probe is suitable for use with both human OGA, as well as the orthologous bacterial counterpart from *Clostridium perfringens, Cp*OGA, and the lysosomal hexosaminidases HexA/B. We structurally characterize *Cp*OGA in complex with a ligand identified from a fragment library screen using this assay. The versatile synthesis procedure could be adapted for making fluorescent probes for the assay of other glycoside hydrolases.



Introduction

Protein O-GlcNAcylation is the dynamic and reversible modification of specific serine and threonine side chains on multitude of nucleocytoplasmic and mitochondrial proteins with a single β-N-acetylglucosamine residue[1]. The attachment of the sugar is catalyzed by the enzyme O-GlcNAc transferase (OGT) and its removal by O-GlcNAcase (OGA). Both OGT and OGA are essential enzymes in mice, with their loss resulting in lethality[2-4]. Protein O-GlcNAcylation is involved in various cellular processes including transcription[5-8], protein stability/degradation[9], and stress response[10-12]; in most these processes, however, the exact role of O-GlcNAcylation is yet to be deciphered. Dysregulation of O-GlcNAc cycling, associated with malfunctions in the production and activity of the processing enzymes, is a feature of pathological conditions such as diabetes, cancer, Alzheimer's disease and cardiovascular disease[1, 9, 13-15].

The structural characterization of bacterial OGAs and more recently human OGA gave insights into the binding mode of O-GlcNAc proteins[16-20]. Schimpl *et al.* showed that glycosylated substrate peptides bind in a conserved groove in a similar conformation and orientation, adopting a "V"-shaped conformation with the residue side chains pointing away from the active site, explaining how a single enzyme can recognize > 1000 different substrates[16]. Inspection of the structures revealed that the peptide backbone forms interactions with OGA, in the -4 through the +3 position surrounding the O-GlcNAc site[16]. Additional hydrophobic interactions are formed by residues in the -1 and -2 position and a surface exposed Tyr^{189} in *Cp*OGA, which is conserved in the human OGA (Tyr^{69})[16, 17]. Studies revealed that mutation of Tyr69 in human OGA leads to a significant reduction in catalytic activity on the substrate analogue 4-methylumbelliferyl-β-D-*N*-acetylglucosaminide (4MU-GlcNAc) [17].

Structural studies of human OGT[21-23] and bacterial orthologues of OGA[24, 25] have allowed the development of inhibitors[26-33] to study the function of these enzymes at the molecular and cellular level. In the case of OGA, potent, cell penetrant inhibitors exist[29-33], and have been used for the identification of OGA as a potential drug target for the treatment of Alzheimer's disease (AD) and cardiovascular disease[34-36]. Hyperphosphorylation of the protein tau resulting in the formation of

neurofibrillary tangles (NFTs) in AD leading to neuronal loss can be slowed by increasing O-GlcNAcylation, which stabilizes tau, preventing its aggregation[34]. This study was performed using mice fed with the OGA inhibitor Thiamet G[30]. A chemically related molecule NButGT[29] was shown to prevent the accumulation of the protein amyloid β, neuroinflammation and memory loss in another mouse model of AD[35]. With respect to cardiovascular disease, inhibition of OGA using NAG thiazolines and NButGT was shown to be protective and attenuate tissue necrosis after ischaemia/reperfusion injury[36, 37]. Furthermore, genetic studies have suggested that deletion of *oga* is perinatally lethal although it is not yet clear whether this essentiality is limited to early development [3, 4]. While potent and selective, the existing inhibitors of OGA possess a carbohydrate scaffold, which has poor drug-like properties with respect to Lipinski's rules. Achieving therapeutic concentrations of these molecules *in vivo* requires the administration of large quantities of the drugs. Molecules with improved stability, pharmacology and penetrability of the blood-brain barrier are desired and the search for these is the subject of ongoing research, exemplified by recent studies describing the use of click chemistry for the rapid generation of potential OGA inhibitor libraries[38] and the use of existing chemical libraries for the discovery of new drug-like inhibitor scaffolds[39].

Fluorescence polarization (FP) is one of many techniques used for the analysis of ligand-protein interactions. The FP technique offers advantages over methods such as surface plasmon resonance (SPR) and biolayer interferometry (BLI) for the measurement of binding affinities of proteins to ligands[40]. It is robust, the quantities of protein/ligands required for an FP assay are comparatively low and neither proteins nor ligands require to be immobilized to a surface. This makes the FP assay completely solution-based and minimizes skewing of binding equilibrium, which could occur by immobilizing a protein, as this may impose constraints on structural changes that could occur upon ligand binding[40, 41]. Also, as a solution based technique, an FP assay is more amenable to high throughput formats. While the OGA activity assay based on the hydrolysis of the fluorescent substrate 4MU-GlcNAc provides a method for the discovery of OGA inhibitors[24, 42], the measurement of inhibition constants for potent inhibitors requires long incubations periods (e.g. ~7 h for *Cp*OGA and

GlcNAcstatin C)[32]. An advantage of an FP based OGA assay is direct measurement of binding rather than loss of activity, reducing assay length. The main limiting factor in the development of an FP assay is the requirement for a tailor-made fluorescent ligand with high affinity to the protein/receptor under study, which is what we report here.

Results and Discussion

Previously, we established a robust synthetic approach to a family of highly potent and selective OGA inhibitors, built on 8-alkylamido 2-phenethyl substituted bicyclic sugar-imidazole scaffold (GlcNAcstatins) [43]. Here we disclose the synthesis of a fluorescently tagged derivative of the potent OGA inhibitor GlcNAcstatin B **1** created by the replacement of the 2-phenethyl group of the parent molecule with aminopropyl spacer labeled with FITC (hereafter referred to as GlcNAcstatin BF **2** (Fig. 1)).

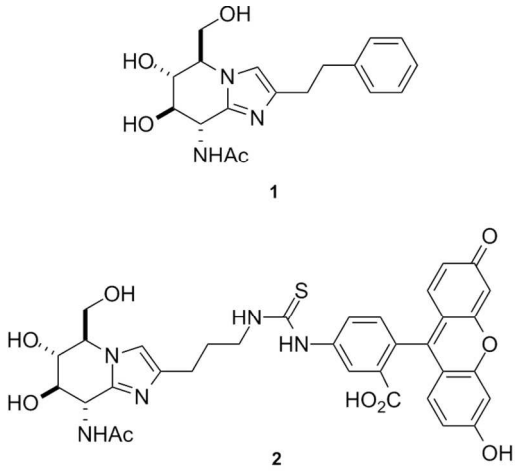
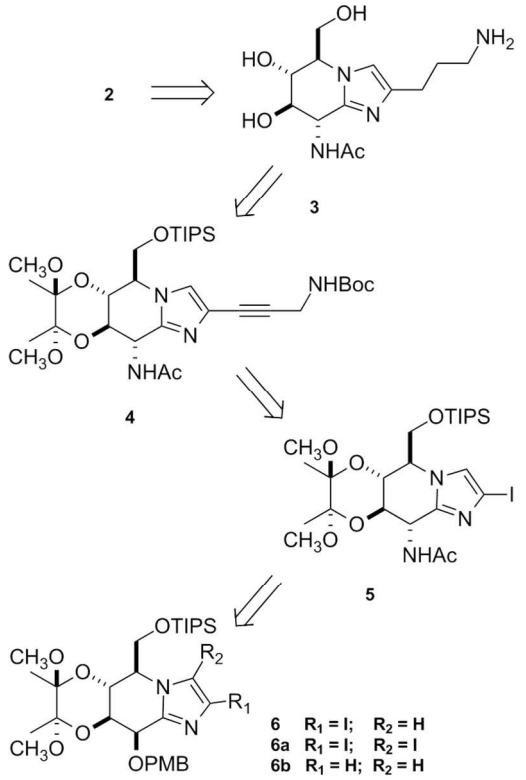


Figure 1. Structures of GlcNAcstatin B (1) and GlcNAcstatin BF (2)

We also report a novel FP displacement assay using GlcNAcstatin BF as a fluorescent tracer to measure the binding affinities of existing inhibitors to the bacterial OGA from *Clostridium perfringens*, *Cp*OGA, as well as human OGA (hOGA).

We initially devised a concise synthetic scheme to access GlcNAcstatin BF **2** (Scheme 1). We scheduled the chemoselective installation of a fluorescent tag onto the derivative **3** bearing the requisite aminopropyl handle that supports the final steps in synthesis of the target compound. To reach **3** we devised a novel synthetic approach for the construction of the 8-alkylamido 2-substituted sugar-imidazole **4** by reversing the order in which the corresponding substituents would be introduced into the target molecule as compared to the published approach[43].



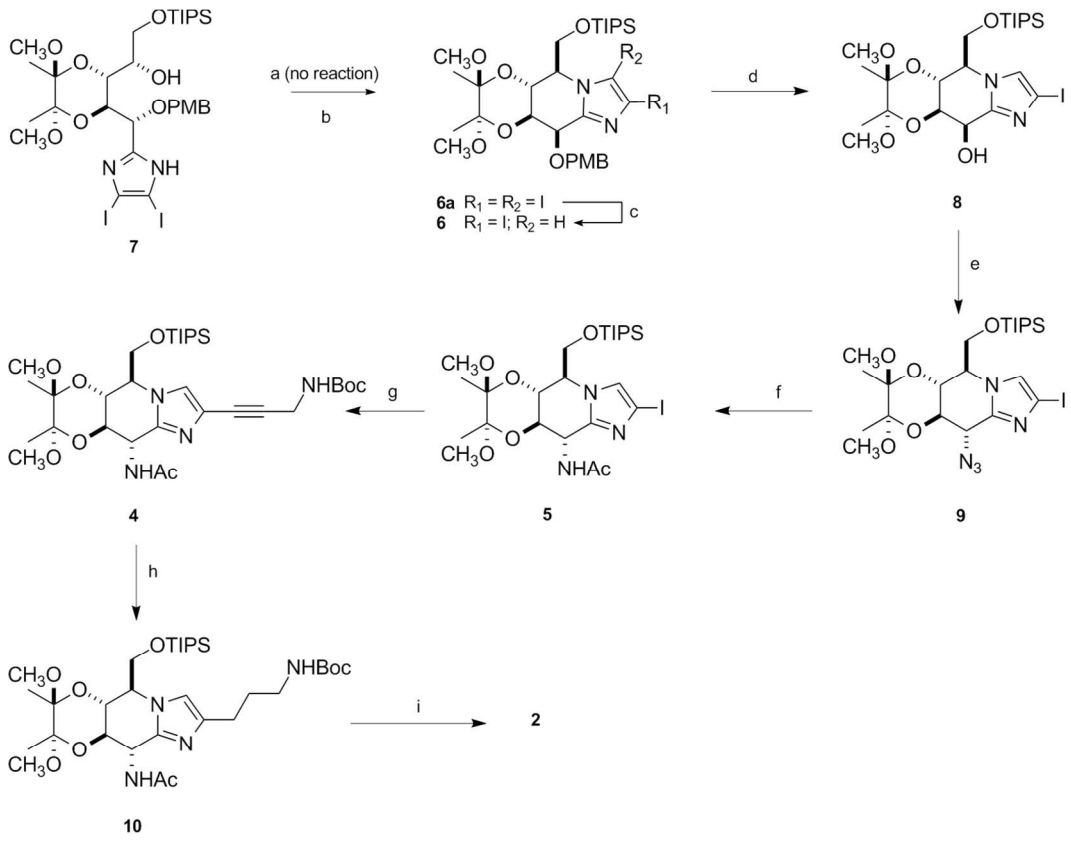
Scheme 1. Retrosynthetic analysis of GlcNAcstatin BF

Accordingly, we planned to transform the known key 2-iodo-mannose-imidazole **6** into GlcNAc-imidazole derivative **5**, which in turn was to be coupled with N-Boc propargyl amine in a Sonogashira reaction giving the intermediate **4**. Although this previously unexplored reverse sequence lacks the flexibility of the previously published approach[43], it sufficed for the case where synthesis of the sole N-acetyl derivative was required.

Previously, the intermediate **6** was prepared by regioselective mono-deiodination of the diiodo-derivative **6a**, which in turn was the product of a challenging bis-iodination (8 eq of NIS, 85 °C, 36 h) of the bicyclic sugar-imidazole **6b**. The latter process was found to be unsuitable for upscaling, resulting in formation of varying amounts of the triply iodinated (*ortho* position in the *p*-methoxybenzyl (PMB) group) product, depending on the quality of NIS used. As an alternative route to intermediate

6, we decided to reinvestigate the cyclization of the linear 4,5-diiodoimidazole **7**, which was previously found to be inert in a triflic anhydride mediated process [43] (Scheme 2).

Scheme 2[a]



[a] Reagents and conditions: (a) Tf_2O, $C_2H_4Cl_2$, 65 °C; (b) PPh_3, DIAD, THF, 65 °C, 1h, 85%; (c) EtMgBr, THF, 0 °C, 1 h, 89%; (d) TFA-H_2O 1:9, DCM, 1h, RT, 85%; (e) DPPA, DBU, Tol, 90 °C, 1h, 95%; (f) PPh_3, THF-H_2O, 65 °C, 3 h, then Ac_2O, DIPEA, 16 RT, 85%; (g) $HCCCH_2NHBoc$; $Pd(PPh_3)_4$, CuI, Et_3N, DMF, 85 °C, 16 h, 53%; (h) H_2, Pd/C, MeOH, 2 h, RT; (i) HCl:TFE 1:50 , 50 °C, 2 h followed by FITC, DIPEA, DMF, RT, 1 h, 65% over two steps.

We assumed that the decreased pK_a of the imidazole moiety in **7** would facilitate intramolecular reactivity under Mitsunobu N-alkylation conditions, giving the bicyclic sugar imidazole **6a** directly. The intermolecular Mitsunobu N-alkylation reaction of imidazoles bearing electron-withdrawing substituents with different alcohols, including secondary, has several precedents in the literature[44, 45]. Furthermore, Mitsunobu type cyclization of structurally similar sugar-derived 1,4-diols was consistently used for the synthesis of C-nucleosides including C2-imidazole and benzimidazole

derivatives[46, 47]. In our hands, reaction of the linear precursor **7** under standard Mitsunobu conditions (PPh_3, DIAD, 65 °C) resulted in formation of the expected known bicyclic product **6a** in excellent yield (Scheme 2). The reaction was shown to be equally applicable to multi-gram scale preparations. Taken together, the novel intramolecular Mitsunobu N-alkylation of **7** in conjunction with the proven highly efficient regioselective mono-deiodination of **6a** represents a streamlined (and more economic) access to the key 2-iodoimidazole **6** that supersedes the triflic anhydride mediated cyclization/bis-iodination sequence of the former approach[43].

We have previously documented the hydrolytic stability of the BDA group in the GlcNAc-imidazoles [43]. Now we reveal that this phenomenon also works for neutral glycoimidazoles to introduce useful orthogonality towards acidolysis between PMB and BDA groups. Indeed, using the standard conditions for the deprotection of the PMB group (aqueous 90% TFA in DCM[48]) we were able to selectively remove the 8-O PMB group in **6** without affecting the DBA[49] and (less surprisingly) the TIPS groups to obtain the monohydroxylic compound **8** in 85 % yield (Scheme 2). This counterintuitive deprotection protocol offers a simplified alternative to the DDQ-mediated PMB removal of the original approach[43].

Using **8** as the substrate, the stereoselective azide introduction with inversion of configuration proceeded smoothly according to the established protocol to give *gluco*-imidazole **9** in a nearly quantitative yield. Next, a one-pot stepwise Staudinger reduction/acetylation sequence applied to **9** gave the crucial 2-iodo-GlcNAc-imidazole **5** in high yield. The Sonogashira coupling of derivative **5** with N-Boc propargyl amine resulted in a fair yield (53%) of the desired advanced intermediate **4**. Overall, we have successfully implemented a novel synthetic approach for the construction of 8-alkylamido-2 substituted sugar–imidazoles. This approach produces comparable efficiency to the original one on all the separate steps but Sonogashira coupling, where somewhat diminished yield was recorded.

The concluding transformation of **4** into the targeted fluorescent tracer **2** was initiated with hydrogenation of the triple bond to give the C-2 aminopropyl equipped derivative **10**. The pivotal one-

pot global removal of the protecting groups in **10** was efficiently achieved with 12 M HCl in trifluoroethanol (1:50) at 55 °C for 2 h to give the fully deprotected compound **3**[50, 51] which was (without purification) treated with a slight excess of FITC in DMF to furnish the fluorescently tagged product **2** (Scheme 2). The reaction mixture was directly purified by reverse phase flash chromatography on a C18 column and the product was re-purified to homogeneity by HPLC to afford the desired GlcNAcstatin BF in 65% yield over the three-step sequence.

We next wanted to assess the use of GlcNAcstatin BF with *Cp*OGA and hOGA in an FP assay. GlcNAcstatin BF bound to *Cp*OGA with a K_d of 0.5 ± 0.2 nM (Fig. 2a) and to hOGA with a K_d of 143 ± 29 nM (Fig. 2b). We then used GlcNAcstatin BF in displacement assays to measure the K_d of the inhibitors GlcNAcstatin B and GlcNAcstatin G. Consistent with previous reports [32, 33], GlcNAcstatin B was a better binder of *Cp*OGA with a K_i of 6.6 nM versus 16.3 nM for GlcNAcstatin G (Fig. 2c). For hOGA, however, the K_i obtained for GlcNAcstatin B was 23.2 nM and for GlcNAcstatin G, it was 21.4 nM (Fig. 2d). The absolute K_i values obtained for these inhibitors are not consistent with previously reported K_i values obtained using the 4MU-NAG based activity assay[32, 33]. In addition to the variability expected from using different methods, this could be due to the differences in the assay buffer and pH used or the different N-terminal truncation of the construct used in this study. Dorfmueller *et al.*, used the McIlvaine buffer system at pH 5.7/7.3 [32, 33], while here we use Tris buffered saline (TBS) at pH 7.5. The K_i values obtained with the FP assay for the binding of *Cp*OGA to GlcNAcstatin G and another potent OGA inhibitor Thiamet G[30] are consistent with those obtained using surface plasmon resonance (SPR) performed using the same assay buffer (Supplementary Fig. S1); the binding affinity of hOGA for Thiamet G using the FP assay is shown in Supplementary Fig. S2. These proof-of-concept experiments establish the fluorescence polarization displacement assay using GlcNAcstatin BF as a convenient method for evaluating affinity of OGA inhibitors.

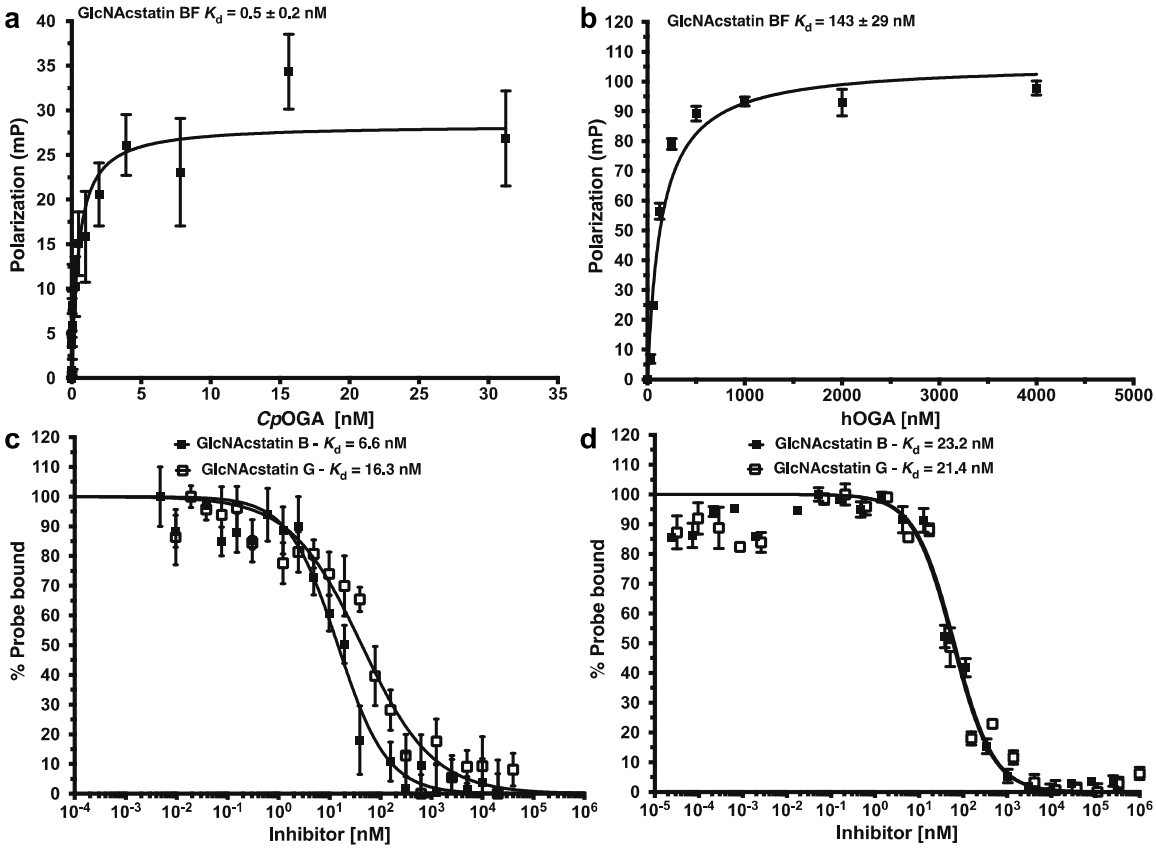


Figure 2. Binding affinity of GlcNAcstatin BF and two inhibitors of the GlcNAcstatin family to *Cp*OGA and hOGA. FP assay showing the binding of GlcNAcstatin BF to **a)** *Cp*OGA and **b)** hOGA. Binding was measured by incubating for 10 min (*Cp*OGA) and 180 min (hOGA) at a fixed concentration of labeled probe ([GBF_{CpOGA}] = 0.5 nM, [GBF_{hOGA}] = 50 nM) with varying concentrations of enzyme. Data points were fitted to a one-site specific-binding equation using Prism (GraphPad). Experiments were performed in triplicate and error bars represent standard error of the mean. Dose-response curves from the fluorescence polarization assay showing the displacement from **c)** *Cp*OGA or **d)** hOGA of a fixed concentration of fluorescent probe by increasing concentrations of GlcNAcstatin B or GlcNAcstatin G. Highest amount of probe bound to enzymes in the absence of inhibitors was set as 100%. Data points were fitted to a four-parameter equation for dose-dependent inhibition using Prism (GraphPad). Experiments were performed in triplicate and error bars represent standard error of the mean. A summary including Hill Slopes can be found in Supplementary Material Table 1.

To expand the applicability of the suggested assay system for the high-throughput screening of custom compound libraries, we decided to evaluate the Maybridge Ro3 1000 fragment library towards *Cp*OGA as a target. This library consists of fragments that adhere to the 'rule of three' (Ro3) (molecular weight ≤ 300 Da, no more than 3 hydrogen bond donors/acceptors, cLogP ≤ 3). The

screen was performed under experimental conditions matching those used in the displacement of GBF by Thiamet G and GlcNAcstatin G (Supplementary material Fig. S3).

The quality of both screens was estimated by calculating Z' value across the controls on a given plate [1], which is a measure of the statistical effect size and therefore the suitability of a high-throughput assay system. In our hands, we consistently saw Z' values for each plate between 0.73 – 0.84 with respect to the standard readings, with a Z' value of 0.66 for readings across all four plates. We chose to identify compounds as hits if they reduced the maximum polarization by ≥ 40%, resulting in a list of 15 initial hits. Subsequently, the initial hits were advanced into the second round of testing, in which we determined the K_i values by measuring the displacement of GlcNAcstatin BF in a dose-dependent manner. Out of the 15 initial hits (F1-F15, Supplementary Material Fig. 3, Supplementary Material Table 2), 8 were confirmed as binders of *Cp*OGA with apparent K_i values between 9 – 150 μM.

Binding of a fragment hit to Tyr^{189} and Asp^{401} blocks GlcNAcstatin BF and substrate binding

To further validate the fragments identified as binders in the HTS screen, we performed macromolecular x-ray crystallography to identify the binding mode of a subset of these hits. We managed to solve the structure of *Cp*OGA in complex with 5-(trifluoromethyl)-2,3-dihydro-1*H*-1,4-diazepine (Fragment F8 hereafter referred to as 5TFD, apparent K_i = 146 μM, Fig. 3A). We collected synchrotron diffraction data to 2.6 Å (Supplementary material Table 3) allowing structure solution by molecular replacement and subsequent refinement (final R_{work}/R_{free} = 0.17/0.22). There was continuous $|F_o|-|F_c|$ electron density for the fragment (Fig. 3B). 5TFD binds proximal to the active site pocket, forming hydrophobic stacking interactions with Tyr^{189} (Fig 3B). Additionally, the secondary amine of 5TFD forms a hydrogen bond with the backbone carbonyl oxygen of Asp^{401} (Fig. 3B). Tyr^{189} has been previously identified to be important for substrate binding, forming hydrophobic interactions with the substrate peptide backbones[16]. When compared to a structure of *Cp*OGA in complex with

GlcNAcstatin F (Fig. 3 C, PDBID: 2XPK [52]), a nM inhibitor of *Cp*OGA designed to mimic the reaction intermediate, 5TFD forms different interactions. However, taking into account the longer linker and presence of the fluorescein group present in GlcNAcstatin BF it is possible for 5TFD to form a steric barrier for proper GBF binding. Thus, the 5TFD fragment bound to Tyr^{189} and Asp^{401}, is able to block binding of GlcNAcstatin BF and glycopeptide substrates.

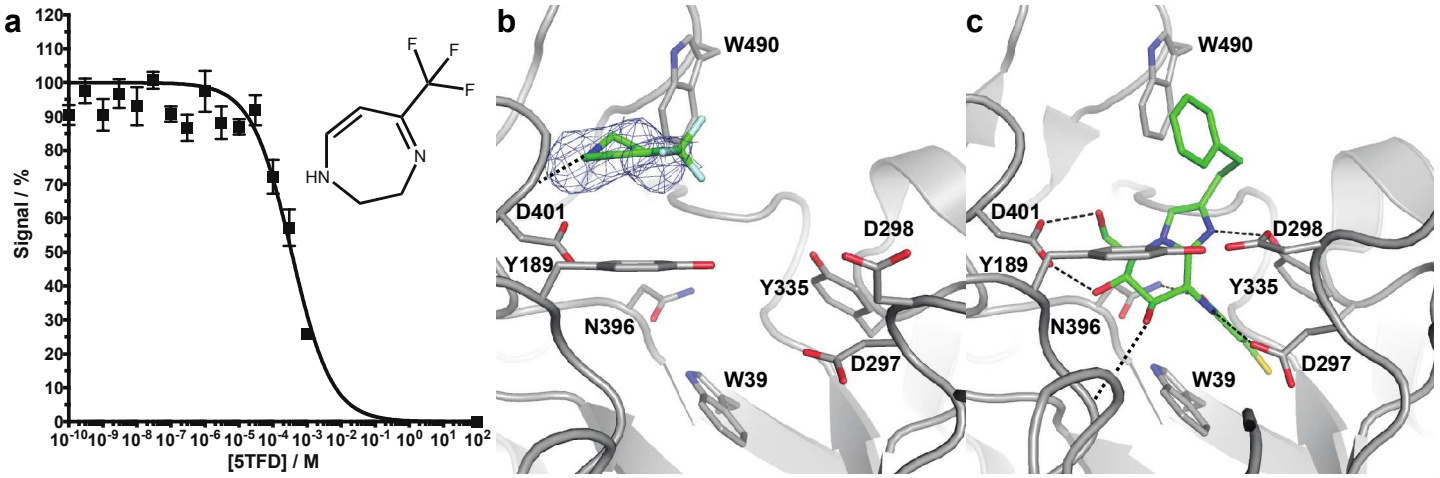


Figure 3. Binding affinity of 5FTD and binding mode of 5FTD and GlcNAcstatin C to CpOGA. a) Dose-response curves from the fluorescence polarization assay showing the displacement from *Cp*OGA of a fixed concentration of fluorescent probe by increasing concentrations of 5FTD. Highest amount of probe bound to enzymes in the absence of inhibitors was set as 100% and a fictive concentration of 10 M set as 0%. Data points were fitted to a four-parameter equation for dose-dependent inhibition using Prism (GraphPad). Experiments were performed in triplicate and error bars represent standard error of the mean. A summary including Hill Slopes can be found in Supplementary Material Table 2. Structure of CpOGA in complex with **b)** 5FTD and **c)** GlcNAcstatin F (PDBID 2XPK [52]). The protein is shown as grey cartoon with active site residues shown as sticks. Ligands are shown as green sticks. Hydrogen bonds between ligands and active site residues are shown as dashed black lines. The unbiased $|F_o|$-$|F_c|$ electron density map for 5FTD is shown as blue mesh contoured at 2 σ.

GlcNAcstatin BF may not only be useful to measure the binding of ligands to OGA, but also to enzymes of related glycoside hydrolase families such as GH20, to which human HexA/B belong. This is because GlcNAcstatin B itself, while potent, is not a selective inhibitor of OGA[32] and also targets HexA/HexB. We were able to determine an apparent K_d of ~ 4 µM (Supplementary Material Figure 4) for GlcNAcstatin BF to HexA/B isolated from bovine kidneys. HexA/HexB have recently been

explored as targets for enzyme enhancement therapy to treat the lysosomal storage diseases Tay-Sachs and Sandhoff disease[53] and an FP based assay offers an alternative to activity assays for the screening of fragments/compounds that bind to these proteins.

Here we have reported the design and synthesis of the novel fluorescent probe GlcNAcstatin BF to support miniaturized OGA high-throughput assays. To synthesize the target compound we established a novel synthetic approach for the construction of the 8-N alkylamide 2-substiuted sugar–imidazoles by reversing the order of the introduction of the respective substituents into the key intermediate as compared to the original approach. As part of the study, we also developed a novel method for the synthesis of bicyclic sugar-imidazoles via intramolecular Mitsunobu N-alkylation of 4,5-diiodoimidazoles to supersede the challenging sequence of the former approach. We successfully applied a simplified counterintuitive acid hydrolysis procedure for the removal of the PMB protection in the presence of cyclic BDA, the product of the unique substrate-induced orthogonality between PMB and DBA groups. Notably, we established a fast and high yielding procedure for the global deprotection of the penultimate synthetic intermediate including removal of the notoriously stable BDA group that constitutes an expedient alternative to the TFA based method previously used for the GlcNAcstatin synthesis. We have shown that GlcNAcstatin BF is suitable for use in fluorescence polarization assays to measure the binding affinity of OGA inhibitors and identify novel binders from a high-throughput screen of a 1000-member fragment library. Finally, we solved the structure of *Cp*OGA in complex with a binder identified in the screen and were able to show that it binds to Tyr189, an important residue for substrate binding.

Materials and Methods

Crystallography and structure solution

*Cp*OGA was concentrated to 40 mg mL^{-1} in 25 mM Tris-HCl pH 8.0, 20 mM NaCl, 0.5 mM TCEP. Sitting-drop vapor diffusion crystallization experiments were performed by mixing drops in a 1:1 ratio of *Cp*OGA and 0.175 M $CdSO_4$, 0.1 M sodium acetate pH 7.5 and needle shaped crystals appeared after 3-4 days. A 5FTD fragment complex was achieved by transferring crystals into a drop containing 10 mM 5FTD in 0.175 M $CdSO_4$, 0.1 M sodium acetate pH 7.5, 1% DMSO for 4 h prior to cryoprotection with 20% glycerol in mother liquor saturated with 5TFD. Diffraction data was collected at the European Synchrotron Radiation Facility (ESRF) on beamline ID30A-3, were processed with XDS[54] and scaled to 2.6 Å using aimless[55]. 5% of total reflections were set aside as an R_{free} test set. Crystals belonged to space group P61 with one molecule per asymmetric unit, a solvent content of 72% and a Matthews coefficient of 4.5. The structures were solved with MOLREP[56], using chain A of PDB 2YDS[16] as a search model. The structure was fully refined using iterative cycles of Refmac5[57] and COOT[58]. Ligand topology was generated using PRODRG[59]. X-ray diffraction data collection and structure refinement statistics can be found in the supplementary material (Supplementary Material Table 3).

Fluorescence Polarization

Experiments were performed in PerkinElmer, black, 384-well plates and millipolarization units measured using a Pherastar FS plate reader (BMG LABTECH) at excitation and emission wavelengths of 485 nm and 520 nm, respectively. For determination of the equilibrium dissociation constant (K_d) of *Cp*OGA and hOGA for GlcNAcstatin BF, 0.5 nM/50 nM of the probe was incubated with a range of concentrations of protein in 25 µL/ 30µL total reaction volume containing 1 x TBS (25 mM Tris, 150 mM NaCl, pH 7.5) buffer and a final concentration of 1-2% DMSO. Reactions were allowed to stand at room temperature for 10 min for *Cp*OGA and 3 h for hOGA, after which

polarization was measured (equilibrium was reached within these time points). Readings were corrected for background emissions from reactions without enzyme and the K_d was determined by fitting a non-linear regression curve with Prism (GraphPad). To avoid receptor depletion, reaction mixtures for competition binding experiments contained 1 nM fluorescent probe for *Cp*OGA and 50 nM for hOGA, 7 nM of *Cp*OGA/250 nM hOGA (receptor) and a range of concentrations of inhibitors in the aforementioned reaction conditions. The largest amount of fluorescent probe bound to the receptors in the absence of competing ligands was set as 100%. IC_{50} values were determined by fitting dose-response curves with Prism (GraphPad) and converted to K_d as outlined elsewhere[60]. All experiments were performed in triplicate.

High-throughput screen of the Maybridge Ro3 1000 fragment library

The Maybridge Ro3 1000 fragment library (Maybridge) screen was performed in black, 384-well plates (PerkinElemer). Displacement of GlcNAcstatin BF was measured by adding 25 µL of 7 nM *Cp*OGA and 1 nM GlcNAcstatin BF in assay buffer (0.1 M Tris-HCl pH 7.4, 150 mM NaCl, 1% DMSO) to assay plates containing 50 nL of a 0.1 M fragment solutions in DMSO, resulting in a final assay concentration of 200 µM. The plates were allowed to stand for 10 minutes in the dark before reading polarization on a Pherastar FS plate reader (BMG Labtech) at excitation and emission wavelengths of 485 nm and 520 nm, respectively. Readings were corrected for background polarization from reactions containing only 1 nM GlcNAcstatin BF in assay buffer and normalised to readings containing 1% DMSO. Fragments displacing GlcNAcstatin BF by ≥ 40% were classified as hits and were advanced for K_i value determination. Competition binding experiments were conducted under the same assay conditions in 0.1 M Tris-HCl pH 7.4, 150 mM NaCl, 2% DMSO and varying concentrations of fragments. K_i values were calculated as described above.

Accession codes

The atomic coordinates and structure factors have been deposited in the Protein Data Bank[61] under the accession code 5OXD.

Supporting information

The Supporting Information, including synthetic procedures and spectral data for all new compounds, methods, figures and tables as well as the compound characterization checklist and the validation report of the deposited structure are available free of charge via the Internet.

Acknowledgements

We thank the European Synchrotron Radiation Facility (ESRF) for beam time on ID30A-3 and assistance. This work was funded by a Wellcome Trust Senior Research Fellowship (WT087590MA) to DMFvA. We would like to hank Dr. M. Schimpl for purifying hOGA.

Author Contributions

The study was conceived by VSB, NS and DMFvA. Synthesis was performed by VSB; protein purification, protein biotinylation and FP assays were performed by NS; KR performed HTS, FP and hit-validation assays and structural biology; cloning was performed by ATF and SPR was performed by TA and IHN. Data were analysed by all authors. The manuscript was written by VSB, NS, KR and DMFvA.

References

[1] Hart, G. W., Slawson, C., Ramirez-Correa, G., and Lagerlof, O. (2011) Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease, *Annu. Rev. Biochem. 80*, 825-858.

[2] Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny, *Proc. Natl. Acad. Sci. U. S. A. 97*, 5735-5739.

[3] Yang, Y. R., Song, M., Lee, H., Jeon, Y., Choi, E. J., Jang, H. J., Moon, H. Y., Byun, H. Y., Kim, E. K., Kim, D. H., Lee, M. N., Koh, A., Ghim, J., Choi, J. H., Lee-Kwon, W., Kim, K. T., Ryu, S. H., and Suh, P. G. (2012) O-GlcNAcase is essential for embryonic development and maintenance of genomic stability, *Aging Cell 11*, 439-448.

[4] Hanover, J., Keembiyehetty, C., Wang, P., Comly, M., Dwyer, N., Krause, M., and Love, D. (2010) Mouse O-Glcnacase Knockout Reveals Roles for O-Glcnac Cycling in Development, Metabolism and Epigenetics, *Glycobiology 20*, 1453-1453.

[5] Ranuncolo, S. M., Ghosh, S., Hanover, J. A., Hart, G. W., and Lewis, B. A. (2012) Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription in vitro and in vivo, *J. Biol. Chem.*

[6] Fujiki, R., Hashiba, W., Sekine, H., Yokoyama, A., Chikanishi, T., Ito, S., Imai, Y., Kim, J., He, H. H., Igarashi, K., Kanno, J., Ohtake, F., Kitagawa, H., Roeder, R. G., Brown, M., and Kato, S. (2011) GlcNAcylation of histone H2B facilitates its monoubiquitination, *Nature 480*, 557-U188.

[7] Chen, Q., Chen, Y., Bian, C., Fujiki, R., and Yu, X. (2013) TET2 promotes histone O-GlcNAcylation during gene transcription, *Nature 493*, 561-564.

[8] Deplus, R., Delatte, B., Schwinn, M. K., Defrance, M., Mendez, J., Murphy, N., Dawson, M. A., Volkmar, M., Putmans, P., Calonne, E., Shih, A. H., Levine, R. L., Bernard, O., Mercher, T., Solary, E., Urh, M., Daniels, D. L., and Fuks, F. (2013) TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS, *EMBO J. 32*, 645-655.

[9] Yang, W. H., Park, S. Y., Nam, H. W., Kim do, H., Kang, J. G., Kang, E. S., Kim, Y. S., Lee, H. C., Kim, K. S., and Cho, J. W. (2008) NFkappaB activation is associated with its O-GlcNAcylation state under hyperglycemic conditions, *Proc. Natl. Acad. Sci. U. S. A. 105*, 17345-17350.

[10] Zachara, N. E., and Hart, G. W. (2004) O-GlcNAc a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress, *Biochim. Biophys. Acta 1673*, 13-28.

[11] Zachara, N. E., O'Donnell, N., Cheung, W. D., Mercer, J. J., Marth, J. D., and Hart, G. W. (2004) Dynamic O-GlcNAc modification of nucleocytoplasmic proteins in response to stress - A survival response of mammalian cells, *Journal of Biological Chemistry 279*, 30133-30142.

[12] Liu, J., Pang, Y., Chang, T., Bounelis, P., Chatham, J. C., and Marchase, R. B. (2006) Increased hexosamine biosynthesis and protein O-GlcNAc levels associated with myocardial protection against calcium paradox and ischemia, *Journal of Molecular and Cellular Cardiology 40*, 303-312.

[13] Erickson, J. R., Pereira, L., Wang, L., Han, G., Ferguson, A., Dao, K., Copeland, R. J., Despa, F., Hart, G. W., Ripplinger, C. M., and Bers, D. M. (2013) Diabetic hyperglycaemia activates CaMKII and arrhythmias by O-linked glycosylation, *Nature 502*, 372-376.

[14] Marsh, S. A., Powell, P. C., Dell'Italia, L. J., and Chatham, J. C. (2013) Cardiac O-GlcNAcylation blunts autophagic signaling in the diabetic heart, *Life Sciences 92*, 648-656.

[15] Yu, Y., Zhang, L., Li, X., Run, X., Liang, Z., Li, Y., Liu, Y., Lee, M. H., Grundke-Iqbal, I., Iqbal, K., Vocadlo, D. J., Liu, F., and Gong, C. X. (2012) Differential effects of an O-GlcNAcase inhibitor on tau phosphorylation, *PLoS ONE 7*, e35277.

[16] Schimpl, M., Borodkin, V. S., Gray, L. J., and van Aalten, D. M. (2012) Synergy of peptide and sugar in O-GlcNAcase substrate recognition, *Chem Biol 19*, 173-178.

[17] Schimpl, M., Schuttelkopf, A. W., Borodkin, V. S., and van Aalten, D. M. (2010) Human OGA binds substrates in a conserved peptide recognition groove, *Biochem J 432*, 1-7.
[18] Li, B., Li, H., Lu, L., and Jiang, J. (2017) Structures of human O-GlcNAcase and its complexes reveal a new substrate recognition mode, *Nat Struct Mol Biol 24*, 362-369.
[19] Roth, C., Chan, S., Offen, W. A., Hemsworth, G. R., Willems, L. I., King, D. T., Varghese, V., Britton, R., Vocadlo, D. J., and Davies, G. J. Structural and functional insight into human O-GlcNAcase.
[20] Elsen, N. L., Patel, S. B., Ford, R. E., Hall, D. L., Hess, F., Kandula, H., Kornienko, M., Reid, J., Selnick, H., Shipman, J. M., Sharma, S., Lumb, K. J., Soisson, S. M., and Klein, D. J. (2017) Insights into activity and inhibition from the crystal structure of human O-GlcNAcase, *Nat Chem Biol*.
[21] Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011) Structure of human O-GlcNAc transferase and its complex with a peptide substrate, *Nature 469*, 564-567.
[22] Lazarus, M. B., Jiang, J., Gloster, T. M., Zandberg, W. F., Whitworth, G. E., Vocadlo, D. J., and Walker, S. (2012) Structural snapshots of the reaction coordinate for O-GlcNAc transferase, *Nat. Chem. Biol. 8*, 966-968.
[23] Schimpl, M., Zheng, X., Borodkin, V. S., Blair, D. E., Ferenbach, A. T., Schuttelkopf, A. W., Navratilova, I., Aristotelous, T., Albarbarawi, O., Robinson, D. A., Macnaughtan, M. A., and van Aalten, D. M. (2012) O-GlcNAc transferase invokes nucleotide sugar pyrophosphate participation in catalysis, *Nat. Chem. Biol. 8*, 969-974.
[24] Rao, F. V., Dorfmueller, H. C., Villa, F., Allwood, M., Eggleston, I. M., and van Aalten, D. M. (2006) Structural insights into the mechanism and inhibition of eukaryotic O-GlcNAc hydrolysis, *EMBO J. 25*, 1569-1578.
[25] Dennis, R. J., Taylor, E. J., Macauley, M. S., Stubbs, K. A., Turkenburg, J. P., Hart, S. J., Black, G. N., Vocadlo, D. J., and Davies, G. J. (2006) Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity, *Nature Structural & Molecular Biology 13*, 365-371.
[26] Gross, B. J., Kraybill, B. C., and Walker, S. (2005) Discovery of O-GlcNAc transferase inhibitors, *J Am Chem Soc 127*, 14588-14589.
[27] Gloster, T. M., Zandberg, W. F., Heinonen, J. E., Shen, D. L., Deng, L., and Vocadlo, D. J. (2011) Hijacking a biosynthetic pathway yields a glycosyltransferase inhibitor within cells, *Nat. Chem. Biol. 7*, 174-181.
[28] Borodkin, V. S., Schimpl, M., Gundogdu, M., Rafie, K., Dorfmueller, H. C., Robinson, D. A., and van Aalten, D. M. (2014) Bisubstrate UDP-peptide conjugates as human O-GlcNAc transferase inhibitors, *Biochem J 457*, 497-502.
[29] Macauley, M. S., Whitworth, G. E., Debowski, A. W., Chin, D., and Vocadlo, D. J. (2005) O-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors, *J. Biol. Chem. 280*, 25313-25322.
[30] Macauley, M. S., Shan, X., Yuzwa, S. A., Gloster, T. M., and Vocadlo, D. J. (2010) Elevation of Global O-GlcNAc in Rodents Using a Selective O-GlcNAcase Inhibitor Does Not Cause Insulin Resistance or Perturb Glucohomeostasis (vol 17, pg 949, 2010), *Chem. Biol. 17*, 1161-1161.
[31] Dorfmueller, H. C., Borodkin, V. S., Schimpl, M., Shepherd, S. M., Shpiro, N. A., and van Aalten, D. M. F. (2006) GlcNAcstatin: A picomolar, selective O-GlcNAcase inhibitor that modulates intracellular O-GlcNAcylation levels, *J Am Chem Soc 128*, 16484-16485.
[32] Dorfmueller, H. C., Borodkin, V. S., Schimpl, M., and van Aalten, D. M. (2009) GlcNAcstatins are nanomolar inhibitors of human O-GlcNAcase inducing cellular hyper-O-GlcNAcylation, *Biochem. J. 420*, 221-227.
[33] Dorfmueller, H. C., Borodkin, V. S., Schimpl, M., Zheng, X., Kime, R., Read, K. D., and van Aalten, D. M. (2010) Cell-penetrant, nanomolar O-GlcNAcase inhibitors selective against lysosomal hexosaminidases, *Chem. Biol. 17*, 1250-1255.
[34] Yuzwa, S. A., Shan, X., Macauley, M. S., Clark, T., Skorobogatko, Y., Vosseller, K., and Vocadlo, D. J. (2012) Increasing O-GlcNAc slows neurodegeneration and stabilizes tau against aggregation, *Nat. Chem. Biol. 8*, 393-399.

[35] Kim, C., Nam, D. W., Park, S. Y., Song, H., Hong, H. S., Boo, J. H., Jung, E. S., Kim, Y., Baek, J. Y., Kim, K. S., Cho, J. W., and Mook-Jung, I. (2013) O-linked beta-N-acetylglucosaminidase inhibitor attenuates beta-amyloid plaque and rescues memory impairment, *Neurobiol. Aging 34*, 275-285.

[36] Laczy, B., Marsh, S. A., Brocks, C. A., Wittmann, I., and Chatham, J. C. (2010) Inhibition of O-GlcNAcase in perfused rat hearts by NAG-thiazolines at the time of reperfusion is cardioprotective in an O-GlcNAc-dependent manner, *Am. J. Physiol. Heart Circ. Physiol. 299*, H1715-1727.

[37] Champattanachai, V., Marchase, R. B., and Chatham, J. C. (2008) Glucosamine protects neonatal cardiomyocytes from ischemia-reperfusion injury via increased protein O-GlcNAc and increased mitochondrial Bcl-2, *Am. J. Physiol. Cell Physiol. 294*, C1509-1520.

[38] Li, T., Guo, L., Zhang, Y., Wang, J., Li, Z., Lin, L., Zhang, Z., Li, L., Lin, J., Zhao, W., Li, J., and Wang, P. G. (2011) Design and synthesis of O-GlcNAcase inhibitors via 'click chemistry' and biological evaluations, *Carbohydr. Res. 346*, 1083-1092.

[39] Dorfmueller, H. C., and van Aalten, D. M. F. (2010) Screening-based discovery of drug-like O-GlcNAcase inhibitor scaffolds, *Febs Letters 584*, 694-700.

[40] Rossi, A. M., and Taylor, C. W. (2011) Analysis of protein-ligand interactions by fluorescence polarization, *Nat. Protocols 6*, 365-387.

[41] Peterson, K. J., Sadowsky, J. D., Scheef, E. A., Pal, S., Kourentzi, K. D., Willson, R. C., Bresnick, E. H., Sheibani, N., and Gellman, S. H. (2008) A fluorescence polarization assay for identifying ligands that bind to vascular endothelial growth factor, *Analytical Biochemistry 378*, 8-14.

[42] (1981) Isolation and further characterization of bovine brain hexosaminidase C, *Biochimica et Biophysica Acta (BBA) - Enzymology 659*, 255 - 266.

[43] Borodkin, V. S., and van Aalten, D. M. F. (2010) An efficient and versatile synthesis of GlcNAcstatins-potent and selective O-GlcNAcase inhibitors built on the tetrahydroimidazo[1,2-a]pyridine scaffold, *Tetrahedron 66*, 7838-7849.

[44] Zolle, I. M., Berger, M. L., Hammerschmidt, F., Hahner, S., Schirbel, A., and Peric-Simov, B. (2008) New selective inhibitors of steroid 11 beta-hydroxylation in the adrenal cortex. Synthesis and structure-activity relationship of potent Etomidate analogues, *J Med Chem 51*, 2244-2253.

[45] Laha, J. K., and Cuny, G. D. (2011) Synthesis of Fused Imidazoles, Pyrroles, and Indoles with a Defined Stereocenter alpha to Nitrogen Utilizing Mitsunobu Alkylation Followed by Palladium-Catalyzed Cyclization, *J Org Chem 76*, 8477-8482.

[46] Guianvarc'h, D., Fourrey, J. L., Dau, M., Guerineau, V., and Benhida, R. (2002) Stereocontrolled synthesis of heterocyclic C-nuclcosides. Protecting group effect and molecular modeling studies, *J Org Chem 67*, 3724-3732.

[47] Yokoyama, M., Tanabe, T., Toyoshima, A., and Togo, H. (1993) SYNTHESIS OF C-NUCLEOSIDES HAVING TYPICAL AROMATIC HETEROCYCLES AS THE BASE MOIETY, *Synthesis-Stuttgart*, 517-520.

[48] Yan, L., and Kahne, D. (1995) P-Methoxybenzyl Ethers as Acid-Labile Protecting Groups in Oligosaccharide Synthesis, *Synlett*, 523-524.

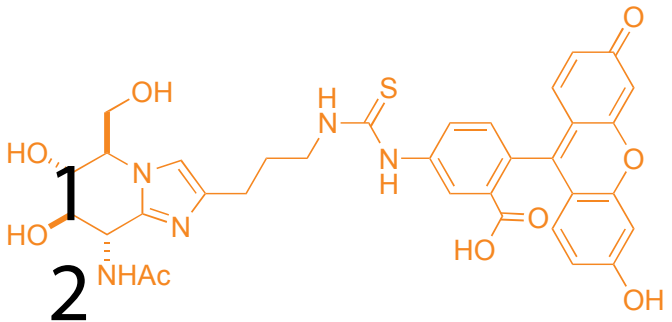
[49] Hense, A., Ley, S. V., Osborn, H. M. I., Owen, D. R., Poisson, J. F., Warriner, S. L., and Wesson, K. E. (1997) Direct preparation of diacetals from 1,2-diketones and their use as 1,2-diol protecting groups, *J Chem Soc Perk T 1*, 2023-2031.

[50] Corey, E. J., Li, W. D., and Reichard, G. A. (1998) A new magnesium-catalyzed doubly diastereoselective anti-aldol reaction leads to a highly efficient process for the total synthesis of lactacystin in quantity, *J Am Chem Soc 120*, 2330-2336.

[51] Palladino, P., and Stetsenko, D. A. (2012) New TFA-Free Cleavage and Final Deprotection in Fmoc Solid-Phase Peptide Synthesis: Dilute HCl in Fluoro Alcohol, *Org Lett 14*, 6346-6349.

[52] Dorfmueller, H. C., Borodkin, V. S., Schimpl, M., Zheng, X., Kime, R., Read, K. D., and van Aalten, D. M. (2010) Cell-penetrant, nanomolar O-GlcNAcase inhibitors selective against lysosomal hexosaminidases, *Chem Biol 17*, 1250-1255.

[53] Tropak, M. B., Blanchard, J. E., Withers, S. G., Brown, E. D., and Mahuran, D. (2007) High-throughput screening for human lysosomal beta-N-acetyl hexosaminidase inhibitors acting as pharmacological chaperones, *Chem. Biol. 14*, 153-164.
[54] Kabsch, W. (2010) Xds, *Acta Crystallogr D Biol Crystallogr 66*, 125-132.
[55] Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr D Biol Crystallogr 67*, 235-242.
[56] Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP, *Acta Crystallographica Section D 66*, 22-25.
[57] <Murshudov, Vagin, Dodson - 1997 - Refinement of macromolecular structures by the maximum-likelihood method.pdf>.
[58] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.
[59] Schuttelkopf, A. W., and van Aalten, D. M. F. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallographica Section D 60*, 1355-1363.
[60] Nikolovska-Coleska, Z., Wang, R., Fang, X., Pan, H., Tomita, Y., Li, P., Roller, P. P., Krajewski, K., Saito, N. G., Stuckey, J. A., and Wang, S. (2004) Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarization, *Anal. Biochem. 332*, 261-273.
[61] Berman, H. M. a. W. J. a. F. Z. a. G. G. a. B. T. N. a. W. H. a. S. I. N. a. B. P. E. (2000) The Protein Data Bank, *Nucleic Acids Research 28*, 235-242.



OGA inhibitors

High affinity fluorescent tracer

Hightroughput OGA

Fluorescent Polarization Assay

ACS Paragon Plus Environment